Approaches to Machine Translation: A Review

[1]John Oladosu, *[2]Adebimpe Esan, *[2]Ibrahim Adeyanju, [3]Benjamin Adegoke, [2]Olatayo Olaniyan and [2]Bolaji Omodunbi
[1]Department of Computer Science and Engineering, LAUTECH, Ogbomoso, Oyo state, Nigeria
[2]Department of Computer Engineering, Federal University Oye-Ekiti, Ekiti state, Nigeria
[3]Department of Computer Engineering, Osun State Polytechnic, Iree, Osun state, Nigeria
jboladosu@lautech.edu.ng, {adebimpe.esan|ibrahim.adeyanju|olatayo.olaniyan|bolaji.omodunbi}@fuoye.edu.ng

Abstract— Translation is the transfer of the meaning of a text from one language to another. It is a means of sharing information across languages and therefore essential for addressing information inequalities. The work of translation was originally carried out by human translators and its limitations led to the development of machine translators. Machine Translation is a subfield of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. There are different approaches to machine translation. This paper reviews the two major approaches (single vs. hybrid) to machine translation and provides critique of existing machine translation systems with their merits and demerits. Several application areas of machine translation and various methods used in evaluating them were also discussed. Our conclusion from the reviewed literatures is that a single approach to machine translation fails to achieve satisfactory performance resulting in lower quality and fluency of the output. On the other hand, a hybrid approach combines the strength of two or more approaches to improve the overall quality and fluency of the translation.

Keywords— language, machine translation, text evaluation

1. Introduction

Language is an efficient medium of communication which expresses the human mind (Abiola, Adetunbi and Oguntimilehin, 2015). There are over 6,800 languages in the world today and this reflects the scope of linguistic diversity. The means of sharing information across languages is known as translation and is essential for addressing information inequalities (Ayegba, Osuagwu and Okechukwu, 2014). Research shows that for translation to be complete, it must have knowledge of the language from which it is translating and profound understanding of the syntax and grammatical features of the two languages involved as well as their vocabularies (Abiola, Adetunbi and Oguntimilehin, 2015). However, early work on translation was carried out by human translators but its limitations (subjective and expensive) led to the development of machine translators (Ayegba, Osuagwu and Okechukwu, 2014).

Machine Translation (MT) can be defined as a subfield of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another (Arnold et al., 1994). It is an area of applied research that draws ideas and techniques from linguistics, computer science, Artificial Intelligence (AI), translation theory and statistics. The goal of machine translation is to develop a system that accurately produces a good translation between human languages (Adeoye, 2012). There are various approaches to machine translation and this paper discusses a chronological review of these approaches.

MT approaches can be divided into two: single approach and hybrid approach. Existing work (Akinwale et al., 2015) reveal that using a single approach to MT fails to achieve satisfactory performance. This is because the approach is sometimes inconsistent, inflexible for large scale applications and uses shallower representation of knowledge resulting in lower quality and fluency of the output. However, a hybrid approach to MT combines the strength of two or more approaches to improve the overall quality of machine translation (Akinwale et al., 2015).

*Corresponding Authors

2. Approaches to Machine Translation

There are various approaches to machine translation. The approaches are classified into two: single and hybrid.

2.1 Single approach

Single approach to machine translation can be referred to as the use of only one method to translate from one natural language to another. The approach includes rule-based, direct-based, corpus-based and knowledge based approaches to machine translation.

2.1.1 Rule-Based Machine Translation (RBMT)

Rule-based approach to machine translation involves the application of morphological, syntactic and semantic rules in the analysis of the source language text and the synthesis of the target-language text (Abiola, Adetunbi and Oguntimilehin, 2015). The approach helps to handle word-order problems and trace parse errors using linguistic knowledge. RBMT can be divided into transfer-based and interlingua based approaches to machine translations respectively.

2.1.2 Direct-based Machine Translation (DBMT)

Direct approach is referred to as the most primitive approach to machine translation. It replaces words in the source language with words in the target language in the same sequence without much linguistic analysis and processing. The bilingual dictionary is the major resource used by this approach (Venkateswara and Muthukumaran, 2013). After word by word translation with the help of a bilingual dictionary, syntactic rearrangement and re-ordering follows. DBMT systems starts with morphological analysis which removes morphological inflections from the words to get the root word from the source language words. Then a bilingual dictionary is looked up to get the target-language words corresponding to the source-language words. This approach is designed for unidirectional translation between one pair of languages (Sharma, 2011).

2.1.3 Corpus-Based Approach

The corpus-based approach is classified into two. They are: example-based machine translations (EBMT) and statistical machine translation (SMT).

i. **Example-based approach:** uses previous translation examples to generate translations for an input provided (Tripathi and Sarkhel, 2010).

ii. **Statistical machine translation (SMT):** is an approach to MT that is characterized by the use of machine learning methods. A learning algorithm is applied to a large body of previously translated text, known as a parallel corpus, and then the learner is able to translate previously unseen sentences. In order to build a functioning SMT system, four problems must be solved which are: translational equivalence model, parameterization, parameter estimation and decoding (Lopez, 2008). In decoding, two models are mainly used: the noisy channel model and the log-linear model.

a. Noisy channel model: Given an *f* sentence, the decoding problem is to find the target sentence *e* which with highest probability could have generated *f*. This best *e* is found using Bayes rule as in equations (i) and (ii):

$$e^* = argmax_e\ P(e/f) \quad \text{(i)}$$

$$e^* = argmax_e\ P(e)P(e/f) \quad \text{(ii)}$$

b. Log-linear model: Recent approaches to SMT extend the noisy channel model to log-linear model as in equation (iii). Log-linear models define relationship between a set of k fixed features h^k(e,d,f) of data and the function P(e,d,f). A feature can be any function $h: E^* \times D^* \times F^* \rightarrow [0, \infty]$ that maps every pair of input and output string to a non-negative value. For example, the number of times a particular word appear (e_i,f_j) in data.

$$P(e,d/f) = \frac{exp \sum_{k=1}^{k} \lambda_k h_k(e,d/f)}{\sum_{e',d': Y(e',d')} exp \sum_{k=1}^{k} \lambda_k h_k(e'd',f)} \quad \text{(iii)}$$

Where k parameters = λ_1^k are called feature weights or model scaling factors. They determine the contribution of a feature to the overall value of *P*(e,d/f). Ideally, each parameter would indicate the pair wise correspondence between the feature and the output probability. A positive value $\lambda_k$should indicate that the feature λ_1^k (e, d, f) correlates with P(e,d/f); a negative value should indicate an inverse correlation; and a value near zero should indicate that the feature is not a useful predictor of *P*(e,d/f). (Lopez, 2008). The log-linear model include the following features: bidirectional rule mapping probabilities, bidirectional lexical rule translation probabilities, target language model, number of rules used, number of target words.

2.1.4 Knowledge-Based Approach

Knowledge-based approach uses extensive semantic and pragmatic knowledge in machine translation. It entails the ability to reason about concepts (Sharma, 2011).

2.2 Hybrid Approach

The approach is a combination of the statistical approach and one or more MT approaches. Mostly, hybrid MT approach is a combination of the statistical method and the rule-based approach. The approach include: word based model, phrase based model, syntax based model and forest based model.

2.2.1 Word-based Models

Word-based models were designed by Brown *et al*(1993) to model the lexical dependencies between single words. Various approaches to word based models also known as alignment models differ in two dimensions which include: the cardinality of the relation between source and target words and the dependency assumptions involved in this mapping. Though, the IBM word-based models were a breakthrough in pioneering the work in SMT, one of their general shortcomings is that they are mainly designed to model the lexical dependencies between single words. They cannot model long-distance re-ordering of words. Hence, phrase-based models were introduced.

2.2.2 Phrase-based Models

Phrase based models were introduced to alleviate the short coming of the word-based models by the introduction of phrases as the basic translation unit (Och, 2003). Phrases can be any substring and they allow local re-orderings, translation of short idioms, or insertions and deletions that are sensitive to local context. They are thus a simple and powerful mechanism for machine translation. Decoding in phrase based methods uses a beam-search approach. Though, phrase-based models help to alleviate shortcomings of the word based models, they still have their shortcomings one of which is the inability to model long distance reordering of source words. This led to the introduction of syntax-based models.

2.2.3 Syntax based model

Syntax is the hierarchical structure of a natural language sentence. Depending on the type of input, syntax-based models can be divided into two broad categories: the string-basedsystems and tree-basedsystems.

i. **String-Based Systems:** These are MT systems whose input is a string to be simultaneously parsed and translated by a synchronous grammar (Galley *et al,* 2006).

In a synchronous CFG the elementary structures are rewrite rules with aligned pairs of right-hand sides as in equation (iv).

$$X \rightarrow \langle \gamma, \alpha, \sim \rangle \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{(iv)}$$

where X is a non-terminal, γ and $\alpha$are both strings of terminals and non-terminals, and, $\sim$ is a one-to-one correspondence between non-terminal occurrences on the source and target side.

ii. **Tree-Based Systems**: They perform translation in two separate steps: parsing and decoding. A parser first parses the source language input into a 1-best tree T, and the decoder then searches for the best derivation(a sequence of translation steps) d*that converts source tree T into a target-language string among all possible derivations D as shown in equation (v):

$$d^* = argmax_{d \in D}\, P(d/T) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(v)}$$

Tree-based systems offer some attractive features. Among these features are: much faster decoding (linear time vs. cubic time, do not require a binary-branching grammar as in string-based models and can have separate grammars for parsing and translation (Huang, Knight and Josi, 2006). Tree-based systems can be sub-divided into three, which are: tree-to-string, string-to-tree and tree-to-tree models respectively. Joshua (Li *et al*, 2009) is an open source toolkit for parsing in syntax-based machine translation. In string-to-tree model, the input is a string and the output is a parse tree (Galley *et a.*, 2006) while tree-to-string model parses a tree at the input and outputs a string. The Tree-to-tree model extracts rules using parse trees from both side(s) of the bi-text (Melamed, 2004). By modeling the syntax of both source and target languages, tree-to-tree model have the potential benefit of providing rules linguistically better motivated (Graehl and Knight, 2004; Melamed, 2004).

However, despite the advantages of tree-based systems, they suffer from a major drawback: they only use the 1-best parse tree to direct the translation, which potentially introduces translation mistakes due to parsing errors (Quirk and Corston-Oliver, 2006). In order to alleviate parsing errors and sparseness of translation rules, tree based system was extended to forest-based MT.

2.2.4 Forest-Based Translation

Forest-based translation is a compromise between the string-based and tree-based methods because it combines the advantages of both methods. Forest based translation encourages faster decoding and alleviates parse errors. Informally, a packed parse forest, or forest in short, is a compact representation of all the derivations (i.e., parse trees) for a given sentence under a context-free grammar (Billot and Lang, 1989; Mi*et al.*, 2008). Forest based machine translation mainly extends the tree-string model to forest to string.

Forest-to-string translation is an extension of the tree-to-string model because it uses a packed parse forest as the input and outputs a string (Mi*et al*, 2008). Forest-to-string models can be described as equation (vi):

$$e^* = argmax_{d \in D(T), T \in F(f)} P(d/T) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(vi)}$$

where f stands for a source string, e stands for a target string, F stands for a forest, D stands for a set of synchronous derivations on a given tree T, and e* stands for the target side yield of a derivation. In order to deal with word order differences in machine translation and to translate differently all the meanings of an ambiguous input in a forest, forest reordering model was proposed by Cmejrek, (2014). He presented a novel extension of a forest-to-string machine translation system with a reordering model as stated in equation (vii):

$$f_{order} = \sum_{0 \le i < j \le n} -\log P_{order}\left(O_{ij} = O_{ij}^{h} | h\right) \ldots\ldots\ldots\ldots \text{(vii)}$$

Research shows that the method provides improvement from 0.6 up to 1.0 point measured by (Ter− Bleu)/2 metric. Figure 1 shows a summary of various approaches to machine translation.

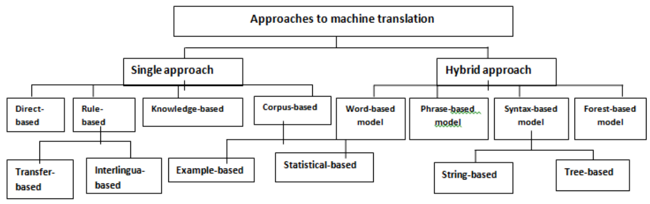

Fig. 1. Approaches to machine translation

3 EVALUATION OF MACHINE TRANSLATION

Traditionally, human judgment is used in evaluating MT systems based on two main criteria: adequacy and fluency. Human judgment of the MT output is expensive and subjective therefore, automatic evaluation measures are a necessity (Adeyanju et al, 2010). There are various methods used in evaluating MT systems. Among them are: BLEU (bilingual evaluation understudy), WER (word error rate), PER (position-independent word error rate) and NIST.

1. **BLEU (bilingual evaluation understudy)**: this is used in automatic calculation of translation quality. It was introduced by Papineni *et al* (2002) to address evaluation problem by comparing the system output against a reference translationof the same source text. The effectiveness of the BLEU metric has been demonstrated by showing that it correlates with human judgment. Two properties of the BLEU metric are: the reliance on higher n-grams and the brevity penalty (BP). Given the precision P_n of n-grams, of size up to N (usually N=4), the length of the test set in words*(c)* and the length of the reference translator *(r)* in words, the equation is computed as in (viii) and (ix):

$$Bleu = Bp.\exp\left(\sum_{n=1}^{4} log P_n\right) \ldots\ldots\ldots\ldots\ldots\ldots \text{(viii)}$$

$$Bp = \min(1, e^{1\,r/c}) \qquad \text{(ix)}$$

2. **WER (word error rate)**: The word error rate is based on the Levenshtein distance. It is computed as the minimum number of substitution, insertion and deletion operations that have to be performed to convert the generated translation into the reference translation. In the case where several reference translations are provided for a source sentence, the minimal distance would be calculated to this set of references as proposed in (Nieben *et al*, 2000). WER measures translation errors. A shortcoming of the WER is that it requires a perfect word order.
3. **PER (position-independent word error rate)**: In order to overcome the shortcoming of the WER, the position independent word was introduced. Words that have no matching counterparts are counted as substitution errors, missing words are deletion and additional words are insertion errors. The PER is a lower bound for WER and measures translation errors.
4. **NIST (National Institute of Standards and Technology):** NIST is a method for evaluating the quality of text which has been translated using machine translation. It is based on the BLEU metric, but with some alterations. Where BLEU simply calculates n-gram precision adding equal weight to each one, NIST calculates how informative a particular n-gram is.

4 EXISTING MACHINE TRANSLATION SYSTEMS

Using a single approach, machine translation has received considerable amount of research attention. Oladosu and Olamoyegun (2012) developed a Yoruba-English language translator for doctor-patient mobile chat. They were motivated by the need to improve rural-urban health care by reducing communication barrier between semi-illiterate patients and highly educated medical personnel who are of different ethnic background. Results show that the application has a high degree of novelty and relevance with about 60% and 80% scores respectively. In addition, Folajinmi and Omonayin, (2012) developed a statistical machine translation (SMT) for English-Yoruba translation. The paper evolved as part of the efforts geared at finding solution to the problems of language barrier in the world and therefore provided tools to tackle the problem of language translation between Yoruba (a Nigerian language) and English language.

Also, Statistical Machine Translation Based Punjabi to English Transliteration System was proposed by Kumar and Kumar (2013). They were motivated by the need to break communication barrier between Punjabi native speakers and those who do not understand Punjabi but English. Results revealed that the system has 97% accuracy when tested on 2000 words and trained on about 15000 words. Moreover, English to Yoruba Machine Translation system using rule-based approach was developed by Eludiora(2014). He was motivated by the need to contribute to knowledge in machine translations by experimenting with an African language.

Agbeyangi, Eludiora and Adenekan, (2015), developed a transfer Rule-Based Machine Translation system. He was motivated by the need to computerize Yoruba language due to its popularity. Results revealed that the system outperforms Google translator which was used as the baseline. Although, these works contributed to efforts geared at finding solution to the problems of language barrier in the world, they are unable to achieve satisfactory performance because they are unsuitable for large scale application.

Hybrid approach to machine translation is an improvement on single approach because it combines the rule based approach with statistical approach. Federmann (2012) described a substitution-based, hybrid machine translation (MT) system from English to German that has been extended with a machine learning component controlling its phrase selection. Also, Arabic-French phrase–based machine translation system was built by Sadiat (2013). She was motivated by the need to build the first Arabic-French phrase–based machine translation system. In addition, Sangeetha, Jothilakshmi and Kumar, (2014) developed a Hybrid MT system for English to Tamil. They were motivated by the need to improve on English to Tamil machine translation. Experimental results show that the proposed approach competes with the machine translation methods reported in the literature and it provides the best translated output in each language with NIST score of 0.8963 and BLEU score of 0.7923. Abiola, Adetunbi and Oguntimilehin (2015), proposed a hybrid approach to English -Yoruba machine translation. They were motivated by the need to improve English to Yoruba machine translation by building a hybrid MT system. Although, these systems improve machine translation greatly in that they are suitable for large scale application, their shortcoming is structural ambiguity in translation.

Moreover, some works were also done to improve hybrid machine translation by combining more than two approaches. Liu *et al* (2010), proposed a forest-based tree-to-tree model that used packed forests. They recorded a significant absolute improvement over state of the art phrase-based systems. Germann (2012) described a syntax-aware phrase-based statistical machine translation model for German-to-English translation. He was motivated by the need to improve German-English translation and experimental results recorded a BLEU score of about 0.15. Mehay and Brew (2012) presented a Combinatory Categorial Grammar (CCG) Syntactic Reordering Models for Phrase-based Machine Translation. They were motivated by the need to improve machine translation from Urdu to English with improvement when compared to a baseline LR system.

A forest-based tree sequence to string model for Chinese-English translation was proposed by Zhang *et al* (2012). They were motivated by the need to integrate the strengths of the forest-based and the tree sequence-based modeling methods. Experimental results show that the system significantly outperforms a baseline system with improvement. Hence, research revealed that a hybrid machine translation approach improves the fluency and adequacy of MT output greatly especially when more than two approaches are combined. Table 1 shows a chronological review of machine translation models.

5 APPLICATION AREAS OF MACHINE TRANSLATION

Machine translation has been integrated in many applications. The areas of application include: e-learning, e-health, commerce, government organizations, production of technical documentation, localization of software, speech translation, information retrieval and information extraction (Hutchins, 2009a; Hutchins, 2009b).

Table 1: Summary of Previous Machine Translation Systems

S/N	Author	MT Approach	Limitation	Merit
1.	Liu et al., (2010)	(Hybrid) forest-based and syntax based models	Structural ambiguity	Suitable for robust application
2.	German, 2012	Syntax based and phrase based models	Structural ambiguity	Suitable for robust application
3.	Zhang et al., (2012)	(Hybrid) forest-based and syntax-based models	Structural ambiguity	Suitable for robust application
4.	Folajinmi and Omonayin, (2012)	(Single) statistical approach	i . Shallow knowledge representation	i. inexpensive ii. suitable for any language pair with enough training data
5.	Oladosu and Olamoyegun, (2012)	(Single) Knowledge-base approach	i. Poor quality and fluency of output. ii.Unsuitable for large scale application	Inexpensive
6.	Federmann (2012),	(Hybrid) Phrase-based	Structural ambiguity	Improved quality and fluency of output
7.	Mehay and brew (2012)	(Hybrid) Phrase based and word based models	Structural ambiguity	Improved quality and fluency of output
8.	Kumar and Kumar (2013)	(Hybrid) Word based model	Unsuitable for long distance reordering of source words.	Handles word order problem
8.	Sadiat, (2013).	(Hybrid) Phrase-based model	Unsuitable for long distance reordering of source words.	Handled local re-ordering and translation of short idioms.
9.	Sangeetha, Jothilakshmi and Kumar (2014)	(Hybrid) Word based model	Unsuitable for long distance reordering of source words.	Handled word order problem
10.	Agbeyangi et al., 2015	(Single) Rule-based	i.Some source text meaning can be lost in the translation ii.Unsuitable for robust application	• Easily handles ambiguities • Easy to build an initial system
11.	Abiola, Adetunbi and Oguntimilehin (2015),	(Hybrid) Word based model	Unsuitable for long distance reordering of source words.	Handles word order problem

1. **E-health**: MT has been used widely in hospitals to break communication barriers between health care givers and patients. An example of this application is MESUDD (Oladosu and Olamoyegun, 2012) which provides real-time language translation for illiterate indigenous patients for interaction with a medical expert system via a small mobile computing device for diagnosis and drug prescription. Oladosu and Emuoyibofarhe (2012) also developed a Yoruba-English Language translator for doctor-patient mobile chat application which help to break communication barrier between doctors and patients in Nigerian hospitals.
2. **Government organizations:** MT systems have been used in governmental organizations for translation of internal documents and assisting administrators in composing texts in non-native languages.
3. **Production of technical documents:** MT systems can be used for production of technical documents. An example is the Logos system used by Ericsson and the Union Bank of Switzerland (Hutchins, 2009a).
4. **Localization of software**: MT can be applied in software localization by making available supporting documentation for new software in many languages.
5. **Speech translation**: Machine translation can be applied in the area of speech translation. Although, there are complexities in speech translation, it can be reduced by restricting communication to relatively narrow domains such as business communication, booking of hotel rooms and negotiating meeting dates.
6. **Information retrieval:** Machine translation finds its application in information retrieval. Information retrieved can be in form of text, images, spoken documents and broadcast stories.
7. **Information extraction:** Many commercial and government-funded international and national organizations have to scrutinize foreign-language documents for information relevant to their activities from commercial and economic to surveillance, intelligence and espionage. Searching can focus on single texts, multilingual collection of texts, selected databases via syndicated feeds or the whole Internet.
8. **E-learning:** MT can be used in Language Learning and tutoring systems for curricular activities.

6 CONCLUSION

Machine translation systems using a single approach fails to achieve satisfactory performance because they are inconsistent and inflexible for large scale application and they give shallower representation of knowledge resulting in lower quality and fluency of the output. Also, hybrid approach to machine translation combines the strength of two or more approaches to improve the overall quality of machine translation. Hence, future researchers should focus more on machine translation using two or more approaches in order to further improve the overall quality of translation. Also, research work should be intensified in mobile based machine translation so as to have remote access to machine translation systems.

REFERENCES

Abiola O.B, Adetunmbi A.O and Oguntimilehin A. (2015). Using hybrid approach for English-to-Yoruba text to text machine translation system (proposed)". *International Journal of Computer Science and Mobile Computing,* 4(8), 308-313.

Abraham, R. C. (1958). *Dictionary of Modern Yoruba*. London: Hodder and Stoughton.

Adeyanju I., Wiratunga N., Lothian R. & Craw S. (2010). Applying Machine Translation Evaluation Techniques to Textual CBR, *Proceedings of Eighteenth International Conference on Case Based Reasoning (ICCBR'10)*, LNCS 6176, 21-35, Springer press, Berlin.

Agbeyangi, A.O., Eludiora, S.I. & Adenekan, O.A. (2015). English to Yorùbá Machine Translation System using Rule-Based Approach. *Journal of Multidisciplinary Engineering Science and Technology*, 2(8).

Akinwale, O. I., Adetunmbi, A. O., Obe, O. O. & Adesuyi, A. T., (2015). Web-Based English to Yoruba Machine Translation *International Journal of Language and Linguistics*. 3(3), 154-159.

Arnold, D., Balkan, L., Meijer, S., Humphreys, R. L., Sadlery, L., (1994). *Machine Translation: An Introductory Guide.* London: NCC Blackwell.

Ayegba, S.F., Osuagwu, O.E. & Okechukwu, N.D. (2014). Machine Translation of Noun Phrases from English to Igala using the Rule-Based Approach".

Billot, S. & Lang, B. (1989). The structure of shared forests in ambiguous parsing: *In Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pp 143- 151

Brown, P.E., Della, V. J., Della, S. A., Mercer, R. L., (1993). The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*, 19(2), pp. 263-311.

Cmejrek, M. (2014). Reordering Model for Forest-to-String Machine Translation. *In proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 227–232

Eludiora, S.I. (2014). *Development of an English to Yoruba Machine Translation System.* Unpublished PhD thesis, Obafemi Awolowo University, Ile-Ife, Osun State, Nigeria

Federmann, C. (2012). Can Machine Learning Algorithms Improve Phrase Selection in Hybrid Machine Translation? *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 113–118.

Folajinmi, Y. & Omonayin, I. (2012). Using Statistical Machine Translation As A language Translation tool forunderstanding Yoruba. *EIE's 2nd International Conference on Computing Energy, Networking, Robotics and Telecommunications*, pp 86-91.

Galley, M., Graehl, J., Knight, K., Marcu, D., DeNeefe, S., WeiWang, & Thayer, I. (2006). Scalable Inference And Training Of Context-Rich Syntactic Translation Models. *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pp. 961–996

Germann, U. (2012). Syntax-aware Phrase-based Statistical Machine Translation: System Description. *In proceedings of the 7th Workshop on Statistical Machine Translation,* pp. 292–297

Graehl, J. & Knight, K. (2004). Training tree transducers. *In Proceedings of HLT-NAACL* (pp. 105–112). Boston: MA,

Huang, L., Knight, K. & Joshi, A. (2006). Statistical syntax-directed translation with extended domain of locality. *In Proceedings of the 7th Conference of the Association for Machine Translation in the Americas*, pp. 223-231

Hutchins, J. (2009a). The development and use of machine translation systems and computer-based translation tools. *International Symposium on Machine Translation and Computer Language Information Processing*, pp. 13-20

Hutchins, J. (2009b). Uses and Applications of Machine Translation. *Presented at Westminster University*.

Kumar, P. & Kumar, V. (2013). Statistical Machine Based Punjabi to English Transliteration System for Proper Nouns. *International Journal of Application or Innovation in Engineering and Management* (IJAIEM). 2(8), pp. 318-321.

Li Z., Callison-Burch, C., Dyer, C., Ganitkevitch, J., Khudanpur, S., Schwartz, L., Thornton, W. N., Weese, J. & Zaidan, O.F., (2009). Joshua: An Open Source Toolkit for Parsing-based Machine Translation. *Proceedings of the 4th EACL Workshop on Statistical Machine Translation*, pp. 135–139

Liu Y., Yajuan, L. & Liu, Q. (2009). Improving Tree-to-Tree Translation with Packed Forests. *In Proceedings of Association of Computational Linguistics,* pp. 84-91

Lopez, A. (2008). Statistical Machine Translation. *Proceedings of Association of Computational Machinery Surveys*, 40(3).

Mehay, D.N. & Brew, C. (2012). CCG Syntactic Reordering Models for Phrase-based Machine Translation. *Proceedings of the 7th Workshop on Statistical Machine Translation*, pp. 210–221

Melamed, I. D., (2004): "Statistical machine translation by parsing". *In Proceedings of the Association for Computational Linguistics (ACL),* pp. 654–661

Mi, H., Liang, H. & Liu, Q. (2008): "Forest-based translation". *In Proceedings of Association of Computational Linguistic,* pp. 192–199

Nießen, S., Och, F.J., Leusch, G. & Ney H. (2000). An evaluation tool for machine translation: Fast evaluation for MT research. *In Proceedings of 2nd International Conference on Language Resources and Evaluation,* pp. 39–46

Och, F.J. (2003). Minimum error rate training in statistical machine translation. *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics,* pp. 160–167

Oladosu, J.B. & Emuoyibofareh, J.O. (2012). A Yoruba-English Language translator for doctor-patient mobile chat application. *International Journal of Computers and Applications*, 34(3), pp. 149-156.

Oladosu, J.B. & Olamoyegun, (2012). MESUDD: a multi-lingua expert system for rural m-healthcare. *International Journal of Electronic Healthcare,* 7(2), pp 141-156.

Papineni, K., Roukos, S., Ward, T., Zhu, W., (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. *In Proceedings of the 40th Annual Meeting on Association for Computational Linguistics* (pp. 311-318). Stroudsburg: PA, USA

Quirk, C. & Corston-Oliver, S. (2006). The impact of parse quality on syntactically-informed statistical machine translation. *In proceedings of Conference on Empirical Methods in Natural Language Processing*, pp. 56-63

Sadiat, F. (2013). Towards a Hybrid Rule-Based and Statistical Arabic-French Machine System. *In proceedings of Recent Advances in natural Language Processing,* pp. 579-583

Sangeetha, J.S., Jothilakshmi, R.N. & Kumar, D. (2014). An Efficient Machine Translation System for English to Indian Languages Using Hybrid Mechanism. *International Journal of Engineering and Technology (IJET),* 6(4), 1909-1919.

Sharma, N. (2011). *English to Indi Statistical Machine Translation System, Unpublished* Masters Thesis, Tharpa University, India.

Venkateswara, P.T. & Muthukumaran. M.G. (2013). Telugu to English Translation using Direct Machine Translation Approach. *International Journal of Science and Engineering Investigations (IJSEI),* vol. 2(12), pp. 25-32.

Zhang, H., Zhang, M., Aiti, H.L., Chew, A. & Tan, L. (2009). Forest-based Tree Sequence to String Translation Model. *In proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP, pp,172-180,Suntec, Singapore.*